

# THE CONTACT PROCESS ON THE COMPLETE GRAPH WITH RANDOM VERTEX-DEPENDENT INFECTION RATES

JONATHON PETERSON

**ABSTRACT.** We study the contact process on the complete graph on  $n$  vertices where the rate at which the infection travels along the edge connecting vertices  $i$  and  $j$  is equal to  $\lambda w_i w_j / n$  for some  $\lambda > 0$ , where  $w_i$  are i.i.d. vertex weights. We show that when  $E[w_1^2] < \infty$  there is a phase transition at  $\lambda_c > 0$  so that for  $\lambda < \lambda_c$  the contact process dies out in logarithmic time, and for  $\lambda > \lambda_c$  the contact process lives for an exponential amount of time. Moreover, we give a formula for  $\lambda_c$  and when  $\lambda > \lambda_c$  we are able to give precise approximations for the probability a given vertex is infected in the quasi-stationary distribution.

Our results are consistent with a non-rigorous mean-field analysis of the model. This is in contrast to some recent results for the contact process on power law random graphs where the mean-field calculations suggested that  $\lambda_c > 0$  when in fact  $\lambda_c = 0$ .

## 1. INTRODUCTION

The contact process is a simple model for the spread of a disease. The standard model of the contact process on a graph  $G = (V, E)$  is described informally as follows. Fix a parameter  $\lambda > 0$  and a set of vertices  $A \subset V$ . At time  $t = 0$  only the vertices in  $A$  are infected. As time progresses, each uninfected vertex  $x$  becomes infected at rate equal to  $\lambda$  times the number of currently infected neighbors, and each infected vertex becomes a healthy (uninfected) vertex at rate 1. More formally, the contact process is a continuous time Markov process  $\eta_t \in \{0, 1\}^V$  with generator

$$\mathcal{L}f(\eta) = \sum_{x \in V} \left( \eta(x) + (1 - \eta(x)) \lambda \sum_{y \sim x} \eta(y) \right) (f(\eta^x) - f(\eta)),$$

where  $f : \{0, 1\}^V \rightarrow \mathbb{R}$  is a bounded function,  $x \sim y$  means that vertices  $x$  and  $y$  are connected by an edge in  $E$ , and  $\eta^x$  is the configuration obtained from  $\eta$  by switching the value of  $\eta(x)$ . That is,

$$\eta^x(z) = \begin{cases} \eta(z) & z \neq x \\ 1 - \eta(z) & z = x. \end{cases} \quad (1)$$

The contact process is also sometimes referred to as the susceptible-infected-susceptible (SIS) epidemic model.

The behavior of the contact process depends on the parameter  $\lambda$ , and as  $\lambda$  increases the infection spreads faster and it takes a longer amount of time for the contact process to die out (i.e., reach the absorbing state of all healthy vertices). It is then natural to ask if there is a critical values of  $\lambda$  for which the contact process exhibits a phase transition. The contact process on the integer lattice  $\mathbb{Z}^d$  has been well studied, and it is known that there is a  $\lambda_c > 0$  such that the contact process started with a single vertex infected dies out with probability one if  $\lambda < \lambda_c$  and survives forever with positive probability if  $\lambda > \lambda_c$  [Lig99]. On any finite graph the

---

*Date:* May 3, 2010.

*2000 Mathematics Subject Classification.* Primary: 60K35; Secondary: 60K37, 82B26, 05C80.

*Key words and phrases.* Contact process, random environment, phase transition.

J. Peterson was partially supported by National Science Foundation grant DMS-0802942.

contact process always eventually dies out, and thus it is not immediately clear how to define a phase transition. However, for the contact process on  $[-n, n]^d \subset \mathbb{Z}^d$  it is known that for  $\lambda < \lambda_c$  (where  $\lambda_c$  is the critical parameter for the contact process on  $\mathbb{Z}^d$ ) the contact process dies out by time  $C \log n$  with high probability, whereas for  $\lambda > \lambda_c$  the contact process survives for time  $\exp\{cn^d\}$  with positive probability [Lig99]. In general, one says that the contact process on a family of finite graphs is sub-critical if the time until the infection dies out is logarithmic in the number of vertices and is super-critical if with positive probability the infection survives for a time that is larger than any polynomial in the number of vertices of the graph. The critical value  $\lambda_c$  then identifies the phase transition of the contact process from sub-critical to super-critical.

Recently, Chatterjee and Durrett [CD09] and Berger, Borgs, Chayes and Saberi [BBCS05] considered the contact process on two different models of power-law random graphs. A power-law random graph is a general term denoting a class of graphs where the distribution of the degree of a typical vertex has tails that decay like  $Ck^{-\alpha-1}$  as  $k \rightarrow \infty$  for some  $\alpha > 1$  and  $C > 0$ . Physicists had previously studied the contact process on power-law random graphs using non-rigorous mean-field calculations and concluded that if  $\alpha > 3$  then there was a critical value  $\lambda_c > 0$  identifying a phase transition [PSV01b, PSV01a]. However, for the two types of power law random graphs studied in [BBCS05] and [CD09] it was shown that in fact  $\lambda_c = 0$  (i.e., the contact process survives for a long time for any  $\lambda > 0$ ). The long time survival of the contact process implies the existence of a quasi-stationary distribution. The mean-field calculations suggest that the average density  $\rho(\lambda)$  of infected sites in the quasi-stationary distribution satisfies  $\rho(\lambda) \sim C\lambda^\beta$  for some  $\beta > 0$  as  $\lambda \rightarrow \lambda_c^+$ . However, upper and lower bounds on  $\rho(\lambda)$  calculated in [CD09] show that the “critical exponent”  $\beta$  must be different from the mean-field predictions.

In this paper, instead of studying the contact process on a random graph, we use a deterministic graph (the complete graph  $K_n$ ) and instead make the rates at which the infections travel along edges to be random. We will choose the random infection rates in a manner that is inspired by the power-law random graph model of Chung and Lu [CL03]. It turns out that in this model the mean field calculations actually turn out to be correct. We are able to obtain a formula for  $\lambda_c$  and approximations for the average density  $\rho(\lambda)$  of infected sites that agree with the mean-field predictions.

**1.1. Description of the Model.** We now turn to a description of the specific process studied in this paper. We will denote the set of vertices of  $K_n$  by  $[n] := \{1, 2, \dots, n\}$ . Let  $\mathbf{w} = \{w_i\}_{i=1}^\infty$  be a sequence of non-negative real numbers, and assign weights  $\mathbf{w}_n = \{w_i\}_{i=1}^n$  to the vertices of  $K_n$ . Then, given the vertex weights, the rate at which infections are transmitted from  $i$  to  $j$  (or  $j$  to  $i$ ) is equal to  $\lambda w_i w_j / n$ . The rate at which infected vertices become healthy is kept constant at 1. Formally, for fixed  $\mathbf{w}_n$ ,  $\lambda > 0$ , and  $A \subset [n]$  we let  $\eta_t$  be the Markov process with  $\eta_0(\cdot) = \mathbf{1}_A(\cdot)$  and generator

$$(\mathcal{L}_{\mathbf{w}_n} f)(\eta) = \sum_{i=1}^n \left( \eta(i) + (1 - \eta(i)) \sum_{j \neq i} \eta(j) \frac{\lambda w_i w_j}{n} \right) (f(\eta^i) - f(\eta)),$$

where  $\eta^i$  is defined as in (1). The law of  $\eta_t$  in this case will be denoted by  $P_{\mathbf{w}}^{n,A}$ . Corresponding expectations will be denoted by  $E_{\mathbf{w}}^{n,A}$ . Often we will be interested in the cases where either all vertices are initially infected or just a single vertex is initially infected. Thus, for notational convenience we will abbreviate  $P_{\mathbf{w}}^{n,[n]}$  and  $P_{\mathbf{w}}^{n,\{i\}}$  by  $P_{\mathbf{w}}^n$  and  $P_{\mathbf{w}}^{n,i}$ , respectively. Corresponding expectations will be denoted similarly.

We will make the following assumption on the vertex weights

**Assumption 1.** *The sequence of vertex weights  $\mathbf{w} = \{w_i\}_{i=1}^\infty$  is an i.i.d. sequence of random variables with common distribution  $\mu$ .*

Expectations with respect to the measure  $\mu$  on the vertex weights will be denoted  $E_\mu$ .

*Remark 1.1.* The essential property that we use is that  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f(w_i) = E_\mu[f(w_1)]$  for certain rational functions  $f(w)$ . Thus, the statements of the theorems and the proofs remain essentially unchanged by letting  $\mathbf{w}$  be an ergodic sequence or even a “nice” deterministic sequence (c.f. [MR98]) with empirical distributions approximating the measure  $\mu$ .

Before stating our main results, we explain briefly the motivation behind this model. One way of constructing the contact process  $\eta_t$  with random edge weights described above is to construct for each  $i$  a Poisson point process  $N_i(t)$  with rate 1 and for each pair  $i \neq j$  a Poisson point process  $N_{i,j}(t)$  with rate  $\lambda w_i w_j / n$ , such that all the Poisson point processes are independent. At each jump time of the process  $N_i(\cdot)$ , if the vertex  $i$  is infected it becomes healthy, and at each jump time of the process  $N_{i,j}(\cdot)$  if exactly one of the vertices  $i$  or  $j$  is infected then the other vertex becomes infected as well.

If we consider the vertices to represent computers or individuals in a network then the jump times of the process  $N_{i,j}(\cdot)$  represent connections made between the respective individuals or computers (such as human contact or e-mail message sent). Thus, if we only keep track of connections formed over a short time period  $[0, \delta]$ , the resulting random graph will have the edge between  $i$  and  $j$  present with probability  $1 - \exp\{-\delta \lambda w_i w_j / n\} \approx \delta \lambda w_i w_j / n$ . Chung and Lu [CL03] studied a model for random graphs where, given a sequence of vertex weights  $w_i$ , the probability that there is an edge connecting vertices  $i$  and  $j$  is proportional to  $w_i w_j / n$ . If the vertex weights  $w_i$  have power-law tails, then the degree distribution of the resulting random graph has power law tails. Thus, if the distribution  $\mu$  on vertex weights has power law tails we might expect the contact process on  $K_n$  described above to be similar to the contact process on a power law random graph. It is therefore somewhat surprising that the behavior is different from the contact process studied on the power-law random graph models in [CD09, BBCS05] and that the mean-field predictions in this case are actually correct.

**1.2. Main Results.** We now turn to the statements of the main results of the paper. We introduce these via the mean-field heuristics which help to explain them. Assume for now that there are only a finite number of weights and that the number of vertices of weight  $x$  is exactly  $\mu(\{x\})n$ . Let  $p_t(\{x\})$  be the probability that a vertex of weight  $x$  is infected at time  $t$ , and let  $N_t(x) = p_t(x)\mu(\{x\})n$  be the number of vertices of weight  $x$  that are infected at time  $t$ . Then, under the mean field assumption that  $\eta_t(i)$  and  $\eta_t(j)$  are independent (which is not true) we obtain that

$$\frac{dN_t(x)}{dt} = -N_t(x) + \sum_y (\mu(\{x\})n - N_t(x))N_t(y) \frac{\lambda xy}{n}.$$

Recalling that  $p_t(x) = N_t(x)/(\mu(\{x\})n)$  we obtain that

$$\begin{aligned} \frac{dp_t(x)}{dt} &= -p_t(x) + \sum_y (1 - p_t(x))p_t(y) \lambda xy \mu(\{y\}) \\ &= -p_t(x) + \int (1 - p_t(x))p_t(y) \lambda xy \mu(dy). \end{aligned} \tag{2}$$

The above mean-field equation should also hold when the distribution of vertex weights is continuous as well.

If the contact process survives for a long time then there should be a quasi-stationary distribution. Thus, we look for a stationary solution to (2). That is, we want to find a function  $p(x)$

such that

$$p(x) = \int (1 - p(x))p(y)\lambda xy \mu(dy) = \int p(y)\lambda xy \mu(dy) - p(x) \int p(y)\lambda xy \mu(dy). \quad (3)$$

Solving for  $p(x)$  we obtain that

$$p(x) = \frac{\sigma x}{1 + \sigma x}, \quad \text{where } \sigma = \lambda \int yp(y) \mu(dy). \quad (4)$$

Substituting the formula for  $p(x)$  on the left into the equation on the right yields the equation

$$1 = \lambda \int \frac{y^2}{1 + \sigma y} \mu(dy). \quad (5)$$

We wish to characterize for what values of  $\lambda$  (depending on the distribution  $\mu$  of vertex weights) there is a  $\sigma > 0$  which solves (5). Note that the right hand side of (5) is decreasing in  $\sigma$ . Thus if there is a solution to (5) it is unique. Moreover, (assuming  $E_\mu w_1 < \infty$ )

$$\lim_{\sigma \rightarrow \infty} \lambda \int \frac{y^2}{1 + \sigma y} \mu(dy) = 0$$

and

$$\lim_{\sigma \rightarrow 0^+} \lambda \int \frac{y^2}{1 + \sigma y} \mu(dy) = \int \lambda y^2 \mu(dy) = \lambda E_\mu(w_1^2).$$

This leads us to the following definitions. Let

$$\lambda_c := \begin{cases} \frac{1}{E_\mu w_1^2} & \text{if } E_\mu w_1^2 < \infty \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

and for  $\lambda > \lambda_c$  define  $\sigma(\lambda)$  as follows.

$$\sigma(\lambda) \text{ is the unique } \sigma > 0 \text{ that solves } 1 = \lambda E_\mu \left[ \frac{w_1^2}{1 + \sigma w_1} \right]. \quad (7)$$

Our main results are a confirmation of the above mean-field heuristics. The first result verifies the existence of a phase transition at  $\lambda_c$ .

**Theorem 1.2.** *Let the vertex weights  $\mathbf{w}$  be i.i.d. with distribution  $\mu$ , and let  $\lambda_c$  be defined as in (6).*

(i) *If  $\lambda < \lambda_c$ , then there exists a constant  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} P_{\mathbf{w}}^n(\eta_{C \log n} = \mathbf{0}) = 1, \quad \mu - a.s.$$

(ii) *If  $\lambda > \lambda_c$ , then there exists a constant  $c > 0$  such that*

$$\lim_{n \rightarrow \infty} P_{\mathbf{w}}^n(\eta_{e^{cn}} \neq \mathbf{0}) = 1, \quad \mu - a.s.$$

Our second main result states that when the contact process is super-critical, the probability of a vertex being infected may be uniformly approximated by the mean-field prediction (4).

**Theorem 1.3.** *Let the vertex weights  $\mathbf{w}$  be i.i.d. with distribution  $\mu$ , and let  $\lambda_c$  and  $\sigma(\lambda)$  be defined as in (6) and (7), respectively. Then, for any  $\lambda > \lambda_c$  and  $\varepsilon > 0$  there exist constants  $C, c > 0$  (depending on  $\varepsilon$ ) so that*

$$\limsup_{n \rightarrow \infty} \sup_{t \in [C \log n, e^{cn}]} \sup_{i \in [n]} \left| P_{\mathbf{w}}^n(\eta_t(i) = 1) - \frac{\sigma(\lambda)w_i}{1 + \sigma(\lambda)w_i} \right| \leq \varepsilon, \quad \mu - a.s.$$

*Remark 1.4.* We are actually able to prove an upper bound on the probability of a vertex being infected that is slightly better than what is implied by the statement of Theorem 1.3. See Proposition 4.3 for a precise statement.

In their study of the contact process on a random graph, Chatterjee and Durrett [CD09] analyzed the contact process at time  $t = e^{\sqrt{n}}$ . The time  $e^{\sqrt{n}}$  is large enough for the contact process to have stabilized but small enough so that with high probability it has not died out. Thus, the distribution of the contact process at time  $t = e^{\sqrt{n}}$  when started with all vertices infected was called the quasi-stationary distribution. As in [CD09], we define  $\rho_n(\lambda)$  to be the expected number of infected vertices at time  $e^{\sqrt{n}}$  when initially all vertices are infected. That is,

$$\rho_n(\lambda) = E_\mu \left[ \frac{1}{n} \sum_{i=1}^n P_{\mathbf{w}}^n(\eta_{e^{\sqrt{n}}}(i) = 1) \right].$$

Then, we obtain the following simple Corollary of Theorem 1.3.

**Corollary 1.5.** *If  $\lambda > \lambda_c$ , then*

$$\rho(\lambda) := \lim_{n \rightarrow \infty} \rho_n(\lambda) = E_\mu \left[ \frac{\sigma(\lambda)w_1}{1 + \sigma(\lambda)w_1} \right].$$

As mentioned above, upper and lower bounds on the critical exponent of  $\rho_n(\lambda)$  were derived in [CD09]. Since

$$E_\mu \left[ \frac{\sigma w_1}{1 + \sigma w_1} \right] \sim E_\mu[w_1]\sigma, \quad \text{as } \sigma \rightarrow 0^+,$$

Corollary 1.5 implies that  $\rho(\lambda)$  and  $\sigma(\lambda)$  have the same critical exponent. The following Proposition allows us to not only identify the critical exponent of  $\sigma(\lambda)$  but also the leading constants.

**Proposition 1.6.** *Let  $\mu(w_1 > x) \sim Cx^{-(\alpha-1)}$  as  $x \rightarrow \infty$  for some  $C > 0$  and  $\alpha > 2$ , and let  $\lambda_c$  and  $\sigma(\lambda)$  be defined as in (6) and (7). Then, as  $\delta \rightarrow 0^+$ ,*

$$\begin{aligned} \sigma(\delta) &\sim \left( \frac{C\pi(\alpha-1)}{\sin(\pi\alpha)} \right)^{1/(3-\alpha)} \delta^{1/(3-\alpha)} && \text{if } \alpha \in (2, 3) \\ \log \sigma(\delta) &\sim \frac{-1}{2C\delta} && \text{if } \alpha = 3 \\ \sigma(\lambda_c + \delta) &\sim \left( \frac{-\sin(\pi\alpha)}{C\lambda_c^2(\alpha-1)\pi} \right)^{1/(\alpha-3)} \delta^{1/(\alpha-3)} && \text{if } \alpha \in (3, 4) \\ \sigma(\lambda_c + \delta) &\sim \frac{1}{3C\lambda_c^2} \frac{\delta}{\log(1/\delta)} && \text{if } \alpha = 4 \\ \sigma(\lambda_c + \delta) &\sim \frac{1}{\lambda_c^2 E_\mu[w_1^3]} \delta && \text{if } \alpha > 4. \end{aligned}$$

Moreover, the conclusion in the case  $\alpha > 4$  also holds under the assumption that  $E_\mu[w_1^3] < \infty$  (i.e., without any assumption on tail asymptotics of  $\mu$ ).

*Remark 1.7.* Since Theorem 1.3 shows that the infection probabilities agree with the mean-field predictions, these critical exponents agree with the mean-field predictions in [PSV01a]. However, the case  $\alpha = 4$  was not considered in [PSV01a].

*Remark 1.8.* Proposition 1.6 states that in the case  $\alpha = 3$ ,  $\sigma(\delta) = \exp\{-\frac{1}{\delta}(\frac{1}{2C} + o(1))\}$ . Riordan [Rio05] has previously observed a similar fast rate of decay for the fraction of vertices in the giant component of percolation on the Barabási-Albert model of power law random graphs (which have  $\alpha = 3$ ).

Before proceeding to the proofs of the above results, we recall two useful facts about the contact process that we will use. Both of these facts are found in [Lig99] and can be proved using the what is called the “graphical representation.”

- **Monotonicity in infection rates.** Increasing the infection rate along any edge only increases the number of vertices infected at any time  $t$ . In particular, if  $\tilde{\mathbf{w}} \leq \mathbf{w}$  in the sense that  $\tilde{w}_i \leq w_i$  for all  $i$ , an easy coupling argument implies that

$$\tilde{\mathbf{w}} \leq \mathbf{w} \Rightarrow P_{\tilde{\mathbf{w}}}^{n,A}(\eta_t \geq \mathbf{1}_B) \leq P_{\mathbf{w}}^{n,A}(\eta_t \geq \mathbf{1}_B), \quad \forall A, B \subset [n]. \quad (8)$$

- **Self-duality.** For any subsets of vertices  $A$  and  $B$ , the probability that an element of  $B$  is infected at time  $t$  when the infection starts from  $A$  is equal to the probability that an element of  $B$  is infected at time  $t$  when the infection starts from  $B$ . That is,

$$P_{\mathbf{w}}^{n,A}(\eta_t(i) = 1, \text{ for some } i \in B) = P_{\mathbf{w}}^{n,B}(\eta_t(i) = 1, \text{ for some } i \in A).$$

An important special case of this is when  $A = \{i\}$  and  $B = [n]$ . In this case, we obtain that

$$P_{\mathbf{w}}^{n,i}(\eta_t \neq \mathbf{0}) = P_{\mathbf{w}}^n(\eta_t(i) = 1). \quad (9)$$

The remainder of the paper is organized as follows. In Section 2 we consider the special case when the distribution  $\mu$  has finite support. In this case we identify the contact process with a multi-dimensional birth-death chain. By analyzing the birth-death chain we then show that when  $\lambda > \lambda_c$  the contact process survives for time  $e^{cn}$  with high probability, and we also obtain a lower bound on the probability a given vertex is infected at any time  $t \leq e^{cn}$ . In Section 3 we extend the results of Section 2 to the general case by approximating  $\mu$  by a measure with finite support. Next in Section 4 we complete the proofs of Theorems 1.2 and 1.3 by proving the complementary results to those obtained in Sections 2 and 3. That is, we show that the contact process dies out with high probability by time  $C \log n$  when  $\lambda < \lambda_c$ , and we obtain upper bounds on the probability a given vertex is infected at any time  $t \geq C \log n$  when  $\lambda > \lambda_c$ . The main technique used in Section 4 is a comparison of the contact process with a related multi-type branching process. Finally, the proof of Proposition 1.6 is given in Section 5.

**Acknowledgement.** Many thanks to Rick Durrett for the many helpful discussions in the process of writing this paper.

## 2. FINITELY MANY TYPES

In this section we will analyze the above model under the assumption that the support of the  $\mu$  on vertex weights  $w_i$  is finite. Let  $\mathbf{W} = (W_1, W_2, \dots, W_m)$  be the possible vertex weights, each occurring with probability  $p_i = \mu(w_1 = W_i)$ . We will classify the vertices according to the weight that they are assigned. That is, a vertex with weight  $W_i$  will be referred to as a *type*  $W_i$  vertex. Since all type  $W_i$  vertices are equivalent, we need only to keep track of the number  $X_i(t)$  of type  $W_i$  vertices that are infected at time  $t$ . That is,

$$X_i(t) := \# \{j : w_j = W_i, \text{ and } \eta_t(j) = 1\}.$$

Note that  $\mathbf{X}(t) = (X_1(t), X_2(t), \dots, X_m(t))$  is an  $m$ -dimensional birth death process. Given the vertex weights  $\mathbf{w}$ , we denote the law of  $\mathbf{X}(\cdot)$  started at  $\mathbf{X}(0) = \mathbf{x}$  by  $P_{\mathbf{w}}^{n,\mathbf{x}}$ . As with the contact process, we will write  $P_{\mathbf{w}}^n$  for the law of  $\mathbf{X}(\cdot)$  when the associated contact process starts with all vertices infected (that is, when  $\sum_{j=1}^m X_j(0) = n$ ).

The main result of this section is the following, which not only shows that the contact process is super-critical when  $\lambda > \lambda_c$ , but also gives a lower bound on the probability of a given vertex to be infected.

**Proposition 2.1.** *Let the distribution  $\mu$  have finite support, and let  $\lambda_c$  and  $\sigma(\lambda)$  be defined as in (6) and (7). Then, if  $\lambda > \lambda_c$  and  $\eta \in (0, \sigma(\lambda))$  there exists a constant  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} P_{\mathbf{w}}^n \left( \inf_{t \leq e^{Cn}} X_i(t) \geq \frac{\eta W_i}{1 + \eta W_i} p_i n, \forall i = 1, 2, \dots, m \right) = 1, \quad \mu - a.s.$$



*Proof.* For any  $\varepsilon > 0$ , let

$$U_n(\varepsilon) := \left\{ \mathbf{x} \in \mathbb{Z}^m : x_i \geq \frac{\varepsilon W_i}{1 + \varepsilon W_i} p_i n, \forall i = 1, 2, \dots, m \right\}.$$

Also, let  $\mathbf{X}[t_1, t_2] := \{\mathbf{X}(t) : t \in [t_1, t_2]\}$  be the trace of the process  $\mathbf{X}(t)$  between times  $t_1$  and  $t_2$ . The idea of the proof is that there exists  $\delta \in (\eta, \sigma(\lambda))$ ,  $\tau > 0$ , and  $C' > 0$  such that  $\mu - a.s.$ ,

$$\inf_{\mathbf{x} \in U_n(\delta)} P_{\mathbf{w}}^{n, \mathbf{x}}(\mathbf{X}(\tau) \in U_n(\delta), \mathbf{X}[0, \tau] \subset U_n(\eta)) \geq 1 - e^{-C'n}, \quad \text{for all } n \text{ large enough.} \quad (10)$$

That is, we can (for  $n$  sufficiently large) uniformly bound from below the probability that starting from a point  $\mathbf{x} \in U_n(\delta)$  the process  $\mathbf{X}(t)$  a short time later is still in  $U_n(\delta)$  and hasn't exited  $U_n(\eta)$ . By dividing  $[0, e^{C'n/2}]$  into  $e^{C'n/2}/\tau$  intervals of length  $\tau$ , (10) implies that,  $\mu - a.s.$ ,

$$\lim_{n \rightarrow \infty} P_{\mathbf{w}}^n(\mathbf{X}[0, e^{C'n/2}] \subset U_n(\eta)) \geq \lim_{n \rightarrow \infty} 1 - e^{-C'n/2}/\tau = 1.$$

This last statement is equivalent to the conclusion of the proposition.

It remains only to prove (10). The monotonicity of the contact process implies that we only need a lower bound on the inner probability in (10) at a single point. That is,

$$\inf_{\mathbf{x} \in U_n(\delta)} P_{\mathbf{w}}^{n, \mathbf{x}}(\mathbf{X}(\tau) \in U_n(\delta), \mathbf{X}[0, \tau] \subset U_n(\eta)) = P_{\mathbf{w}}^{n, \mathbf{x}^\delta}(\mathbf{X}(\tau) \in U_n(\delta), \mathbf{X}[0, \tau] \subset U_n(\eta)),$$

where

$$\mathbf{x}^\delta = (x_1^\delta, x_2^\delta, \dots, x_m^\delta), \quad x_i^\delta := \left\lceil \frac{\delta W_i}{1 + \delta W_i} p_i n \right\rceil.$$

For any  $0 < \delta < \varepsilon$  let

$$B_n(\delta, \varepsilon) := \left\{ \mathbf{x} \in \mathbb{Z}^m : \frac{\delta W_i}{1 + \delta W_i} \leq \frac{x_i}{p_i n} \leq \frac{\varepsilon W_i}{1 + \varepsilon W_i}, i = 1, 2, \dots, m \right\}.$$

Note that  $B_n(\delta, \varepsilon) \subset U_n(\delta)$ . Then, to prove (10) it is enough to show that for some  $\eta < \delta < \varepsilon < \sigma(\lambda)$ , and constants  $\tau, C' > 0$ ,  $\mu - a.s.$ ,

$$P_{\mathbf{w}}^{n, \mathbf{x}^\delta}(\mathbf{X}(\tau) \in U_n(\delta), \mathbf{X}[0, \tau] \subset B_n(\eta, \varepsilon)) \geq 1 - e^{-C'n}, \quad \text{for all } n \text{ large enough.} \quad (11)$$

The idea behind proving (11) is that by choosing  $\delta$  close enough to  $\eta < \sigma(\lambda)$ , the drift of the process  $\mathbf{X}(t)$  is increasing in all coordinates in a neighborhood of  $\mathbf{x}^\delta$  (with high probability). Then, the process  $\mathbf{X}(t)$  starting at  $\mathbf{x}^\delta$  will with high probability increase in all coordinates by time  $\tau$  and by choosing  $\tau > 0$  small enough it will also not have exited  $B_n(\eta, \varepsilon)$ .

As mentioned above,  $\mathbf{X}(t)$  is an  $m$ -dimensional birth death process (given the vertex weights), but in order to describe the jump rates we need some additional notation. Let

$$N_{i,n} := \#\{j \leq n : w_j = W_i\}, \quad \text{and} \quad \hat{p}_{i,n} := \frac{N_{i,n}}{n}$$

be the number and proportion, respectively, of type  $W_i$  vertices. Note that the law of large numbers implies that

$$\lim_{n \rightarrow \infty} \hat{p}_{i,n} = p_i, \quad \mu - a.s.$$

Let  $\{\mathbf{e}_i\}_{i=1}^m$  be the canonical basis for  $\mathbb{Z}^m$ . If  $\mathbf{X}(t) = \mathbf{x} = (x_1, x_2, \dots, x_m)$ , then  $\mathbf{X}(t)$  increases in the  $i$ th coordinate to  $\mathbf{x} + \mathbf{e}_i$  at rate  $q_i^+(\mathbf{x})$  and decreases in the  $i$ th coordinate to  $\mathbf{X}(t) - \mathbf{e}_i$  at rate  $q_i^-(\mathbf{x})$ , where

$$q_i^+(\mathbf{x}) = (N_{i,n} - x_i) \sum_{j=1}^m x_j \frac{\lambda W_i W_j}{n}, \quad \text{and} \quad q_i^-(\mathbf{x}) = x_i. \quad (12)$$

The following Lemma allows us to bound the jump rates in some  $B_n(\eta, \varepsilon)$ .

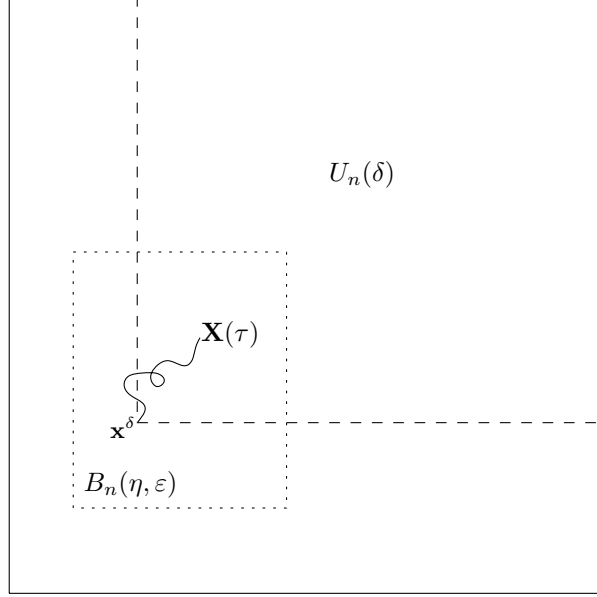


FIGURE 1. An example of the event in (11). The key is to choose  $\eta < \delta < \varepsilon$  so that the drift of the process  $\mathbf{X}(t)$  is increasing in all coordinates inside of  $B_n(\eta, \varepsilon)$ .

**Lemma 2.2.** *For any  $\eta < \sigma(\lambda)$ , there exists an  $\varepsilon \in (\eta, \sigma(\lambda))$  and positive constants  $\{\alpha_i^+, \alpha_i^-\}_{i=1}^m$  such that  $\mu - a.s.$ , the event*

$$\bigcap_{i=1}^m \bigcap_{\mathbf{x} \in B_n(\eta, \varepsilon)} \{q_i^-(\mathbf{x}) \leq \alpha_i^- n < \alpha_i^+ n \leq q_i^+(\mathbf{x})\}$$

*occurs for all  $n$  large enough.*

Postponing for now the proof of Lemma 2.2 we continue with the proof of (11). Let  $\tilde{\mathbf{X}}(t) = (\tilde{X}_1(t), \dots, \tilde{X}_m(t))$ , where the  $\tilde{X}_i(t)$  are independent continuous time simple random walks that increase by 1 at rate  $\alpha_i^+ n$  and decrease by 1 at rate  $\alpha_i^- n$ . Then (2.2) implies that (for  $n$  large enough)  $\mathbf{X}(t)$  stochastically dominates  $\tilde{\mathbf{X}}(t)$  while inside of  $B_n(\eta, \varepsilon)$ . Let  $\tilde{P}^{\mathbf{x}}$  be the law of  $\tilde{\mathbf{X}}(t)$  started at  $\tilde{\mathbf{X}}(0) = \mathbf{x}$ . Since  $\tilde{\mathbf{X}}(t)$  is a continuous time simple random walk with positive drift in all coordinates and jump rates proportional to  $n$ , standard large deviation estimates yield the following Lemma whose proof we omit.

**Lemma 2.3.** *Let  $\alpha_i^- < \alpha_i^+$  for  $i = 1, 2, \dots, m$  be given. Then, there exists a constant  $C'' > 0$  (depending only on the  $\alpha_i^\pm$ ) such that for any  $\delta > 0$ ,*

$$\tilde{P}^{\mathbf{x}^\delta} \left( \tilde{\mathbf{X}}(t) \notin U_n(\delta) \right) \leq m e^{-C'' t n}, \quad \forall t \geq 0.$$

On the event  $\bigcap_{i=1}^m \bigcap_{\mathbf{x} \in B_n(\eta, \varepsilon)} \{q_i^-(\mathbf{x}) \leq \alpha_i^- n < \alpha_i^+ n \leq q_i^+(\mathbf{x})\}$ , since  $\mathbf{X}(t)$  stochastically dominates  $\tilde{\mathbf{X}}(t)$  inside of  $B_n(\eta, \varepsilon)$  we have

$$\tilde{P}^{\mathbf{x}^\delta} \left( \tilde{\mathbf{X}}(\tau) \in U_n(\delta) \right) \leq P_{\mathbf{w}}^{n, \mathbf{x}^\delta} (\mathbf{X}(\tau) \in U_n(\delta), \mathbf{X}[0, \tau] \subset B_n(\eta, \varepsilon)) + P_{\mathbf{w}}^{n, \mathbf{x}^\delta} (\mathbf{X}[0, \tau] \not\subset B_n(\eta, \varepsilon)) \quad (13)$$

Now, the total jump rate of the process  $\mathbf{X}(t)$  at any point  $\mathbf{x}$  is always bounded by  $Cn$  for some constant  $C$  (depending on  $\lambda$  and the  $W_j$ ). Then, since the distance from  $\mathbf{x}^\delta$  to the complement of  $B_n(\eta, \varepsilon)$  is at least  $cn$  for some constant  $c$  (depending on  $\eta, \delta, \varepsilon$  and the  $W_j$ ) by choosing



$\tau > 0$  sufficiently small we obtain that for some  $C_0 > 0$ ,

$$P_{\mathbf{w}}^{n, \mathbf{x}^\delta}(\mathbf{X}[0, \tau] \not\subseteq B_n(\eta, \varepsilon)) \leq e^{-C_0 n}.$$

Thus, applying Lemmas 2.2 and 2.3 to (13) we obtain that  $\mu - a.s.$ ,

$$P_{\mathbf{w}}^{n, \mathbf{x}^\delta}(\mathbf{X}(\tau) \in U_n(\delta), \mathbf{X}[0, \tau] \subset B_n(\eta, \varepsilon)) \geq 1 - me^{-C''\tau n} - e^{-C_0 n} \geq 1 - e^{-C'n},$$

for all large  $n$  if  $C' < \min\{C'', C_0\}$ . This completes the proof of (11) and thus also (modulo the proof of Lemma 2.2) the proof of Proposition 2.1.  $\square$

We now return to the proof of Lemma 2.2.

*Proof of Lemma 2.2.* We ultimately wish to control  $q_i^+(\mathbf{x})$  and  $q_i^-(\mathbf{x})$  inside of some  $B_n(\eta, \varepsilon)$ , but fluctuations in the number of type  $W_i$  vertices force us to deal first with a slightly different set that depends on the actual number of type  $W_i$  vertices. For any  $0 < \eta < \varepsilon$  define

$$\widehat{B}_n(\eta, \varepsilon) := \left\{ \mathbf{x} \in \mathbb{Z}^m : \frac{\eta W_i}{1 + \eta W_i} \leq \frac{x_i}{N_{i,n}} \leq \frac{\varepsilon W_i}{1 + \varepsilon W_i}, i = 1, 2, \dots, m \right\}.$$

Note that the region  $\widehat{B}_n(\eta, \varepsilon)$  depends on what the actual vertex weights are, whereas the region  $B_n(\eta, \varepsilon)$  only depends on the given parameters. However, since  $N_{i,n}/n \rightarrow p_i$ ,  $\mu - a.s.$  we obtain that for any  $\eta' < \eta < \varepsilon < \varepsilon'$

$$B_n(\eta, \varepsilon) \subset \widehat{B}_n(\eta', \varepsilon'), \quad \text{for all } n \text{ large enough, } \mu - a.s. \quad (14)$$

Recall the formulas for the jump rates  $q_i^+(\mathbf{x})$  and  $q_i^-(\mathbf{x})$ . Then, for any  $\mathbf{x} \in \widehat{B}_n(\eta, \varepsilon)$ ,

$$\begin{aligned} q_i^+(\mathbf{x}) &\geq (N_{i,n} - \frac{\varepsilon W_i}{1 + \varepsilon W_i} N_{i,n}) \sum_{j=1}^m \frac{\eta W_j}{1 + \eta W_j} N_{j,n} \frac{\lambda W_i W_j}{n} \\ &= \lambda N_{i,n} \frac{W_i}{1 + \varepsilon W_i} \sum_{j=1}^m \frac{\eta W_j^2}{1 + \eta W_j} \frac{N_{j,n}}{n} \\ &= \lambda \hat{p}_{i,n} \frac{W_i}{1 + \varepsilon W_i} \sum_{j=1}^m \hat{p}_{j,n} \frac{\eta W_j^2}{1 + \eta W_j} n \\ &=: \hat{\theta}_i^+(\eta, \varepsilon, n)n \end{aligned}$$

and

$$q_i^-(\mathbf{x}) \leq \frac{\varepsilon W_i}{1 + \varepsilon W_i} N_{i,n} = \frac{\varepsilon W_i}{1 + \varepsilon W_i} \hat{p}_{i,n} n =: \hat{\theta}_i^-(\eta, \varepsilon, n)n$$

Again, since  $\hat{p}_{i,n} \rightarrow p_i$  we obtain

$$\lim_{n \rightarrow \infty} \hat{\theta}_i^+(\eta, \varepsilon, n) = \lambda p_i \frac{W_i}{1 + \varepsilon W_i} \sum_{j=1}^m p_j \frac{\eta W_j^2}{1 + \eta W_j} =: \theta_i^+(\eta, \varepsilon), \quad \mu - a.s. \quad (15)$$

and

$$\lim_{n \rightarrow \infty} \hat{\theta}_i^-(\eta, \varepsilon, n) = p_i \frac{\varepsilon W_i}{1 + \varepsilon W_i} =: \theta_i^-(\eta, \varepsilon), \quad \mu - a.s. \quad (16)$$

Note that  $\theta_i^+ = (1 + \Delta)\theta_i^-$  where

$$\Delta = \Delta(\eta, \varepsilon) = \frac{\lambda \eta}{\varepsilon} \sum_{j=1}^m \frac{p_j W_j^2}{1 + \eta W_j} - 1 = \frac{\lambda \eta}{\varepsilon} E_\mu \left[ \frac{w_1^2}{1 + \eta w_1} \right] - 1.$$

Now, the definition of  $\sigma(\lambda)$  implies that  $\Delta(\eta, \eta) = \lambda E_\mu[w_1^2/(1+\eta w_1)] - 1 > 0$  for any  $\eta < \sigma(\lambda)$ . Therefore, we can choose an  $\eta' < \eta < \varepsilon' < \sigma(\lambda)$  such that  $\Delta(\eta', \varepsilon') > 0$ . Fix such a  $\eta'$  and  $\varepsilon'$  and choose  $\alpha_i^+$  and  $\alpha_i^-$  so that

$$\theta_i^-(\eta', \varepsilon') < \alpha_i^- < \alpha_i^+ < \theta_i^+(\eta', \varepsilon').$$

Then, (15) and (16) imply that

$$\hat{\theta}_i^-(\eta', \varepsilon', n) \leq \alpha_i^- < \alpha_i^+ \leq \hat{\theta}_i^+(\eta', \varepsilon', n), \quad \text{for all } n \text{ large enough, } \mu - a.s. \quad (17)$$

Now choose  $\varepsilon \in (\eta, \varepsilon')$ . Applying (14) and (17) completes the proof of Lemma 2.2  $\square$

We end this section with the following Corollary.

**Corollary 2.4.** *Let the distribution  $\mu$  have finite support. Then, for any  $\lambda > \lambda_c$  and  $\eta < \sigma(\lambda)$ , there exists a constant  $c > 0$  such that*

$$\liminf_{n \rightarrow \infty} \inf_{t \leq e^{cn}} \inf_{i \in [n]} \left( P_{\mathbf{w}}^n(\eta_t(i) = 1) - \frac{\eta w_i}{1 + \eta w_i} \right) \geq 0, \quad \mu - a.s.$$

*Proof.* For  $\lambda > \lambda_c$  and  $\eta < \sigma(\lambda)$  fixed, chose  $c > 0$  as in the statement of Proposition 2.1. Suppose that vertex  $i$  is a type  $W_j$  vertex. That is  $w_i = W_j$ . Then, since there are  $N_{j,n}$  vertices of type  $W_j$ ,

$$E_{\mathbf{w}}^n[X_j(t)] = N_{j,n} P_{\mathbf{w}}^n(\eta_t(i) = 1).$$

Therefore,

$$\begin{aligned} \inf_{t \leq e^{cn}} \inf_{i \in [n]} \left( P_{\mathbf{w}}^n(\eta_t(i) = 1) - \frac{\eta w_i}{1 + \eta w_i} \right) &= \inf_{t \leq e^{cn}} \inf_{j \leq m} \left( \frac{1}{N_{j,n}} E_{\mathbf{w}}^n[X_j(t)] - \frac{\eta W_j}{1 + \eta W_j} \right) \\ &\geq \inf_{j \leq m} \left( \frac{1}{N_{j,n}} E_{\mathbf{w}}^n \left[ \inf_{t \leq e^{cn}} X_j(t) \right] - \frac{\eta W_j}{1 + \eta W_j} \right). \end{aligned} \quad (18)$$

Now, Proposition 2.1 and the fact that  $N_{j,n} \sim p_j n$  as  $n \rightarrow \infty$  imply that

$$\liminf_{n \rightarrow \infty} \frac{1}{N_{j,n}} E_{\mathbf{w}}^n \left[ \inf_{t \leq e^{cn}} X_j(t) \right] \geq \frac{\eta W_j}{1 + \eta W_j}, \quad \mu - a.s. \quad (19)$$

Combining (18) and (19) completes the proof.  $\square$

### 3. INFINITELY MANY TYPES

In this section we will extend the results of the previous section to the case where  $\mu$  does not have finite support. The first main result is a proof of part (ii) of Theorem 1.2.

*Proof of Theorem 1.2, part (ii).* For any integer  $m \geq 1$  let

$$\kappa_m(x) := \frac{\lfloor xm \rfloor}{m} \wedge m.$$

Note that  $\kappa_m(x) \leq x$  and  $\lim_{m \rightarrow \infty} \kappa_m(x) = x$  for all  $x \in \mathbb{R}$ . Given the sequence of vertex weights  $\mathbf{w} = \{w_i\}_{i=1}^\infty$ , let

$$\mathbf{w}^{(m)} := \{\kappa_m(w_i)\}_{i=1}^\infty.$$

The sequence vertex weights  $\mathbf{w}^{(m)}$  takes on only finitely many values, and thus we may apply the results from the previous section. For any vertex distribution  $\mu$  and  $m \geq 1$ , let  $\lambda_c^{(m)}$  and  $\sigma^{(m)}(\lambda)$  be the analogues of  $\lambda_c$  and  $\sigma(\lambda)$  for the modified vertex weights  $\mathbf{w}^{(m)}$ . That is,

$$\lambda_c^{(m)} := (E_\mu[\kappa_m(w_1)^2])^{-1},$$

and for  $\lambda > \lambda_c^{(m)}$

$$\sigma^{(m)}(\lambda) \text{ is the unique } \sigma > 0 \text{ that solves} \quad 1 = \lambda E_\mu \left[ \frac{\kappa_m(w_1)^2}{1 + \sigma \kappa_m(w_1)} \right].$$

The monotone convergence theorem implies that  $\lambda_c^{(m)} \searrow \lambda_c$  and  $\sigma^{(m)}(\lambda) \nearrow \sigma(\lambda)$  as  $m \rightarrow \infty$ .

If  $\lambda > \lambda_c$ , then there exists some  $m$  large enough so that  $\lambda > \lambda_c^{(m)}$ . For this  $m$  fixed, we may apply Proposition 2.1 to the modified vertex weights  $\mathbf{w}^{(m)}$  to obtain that there exists a  $C > 0$  such that

$$\lim_{n \rightarrow \infty} P_{\mathbf{w}^{(m)}}^n(\eta_{e^{Cn}} \neq \mathbf{0}) = 1, \quad \mu - a.s. \quad (20)$$

Since  $\mathbf{w}^{(m)} \leq \mathbf{w}$ , the monotonicity of the contact process (8) implies that  $P_{\mathbf{w}}^n(\eta_t \neq \mathbf{0}) \geq P_{\mathbf{w}^{(m)}}^n(\eta_t \neq \mathbf{0})$ , and thus (20) holds with  $\mathbf{w}$  in place of  $\mathbf{w}^{(m)}$ .  $\square$

We can also extend Corollary 2.4 to the general case. This proves the lower bound needed for the proof of Theorem 1.3.

**Corollary 3.1.** *Let  $\lambda > \lambda_c$ . Then, for any  $\varepsilon > 0$  there exists a constant  $c > 0$  so that*

$$\liminf_{n \rightarrow \infty} \inf_{t \leq e^{cn}} \inf_{i \in [n]} \left( P_{\mathbf{w}}^n(\eta_t(i) = 1) - \frac{\sigma(\lambda)w_i}{1 + \sigma(\lambda)w_i} \right) \geq -\varepsilon, \quad \mu - a.s.$$

*Proof.* First note that an exercise in calculus shows that

$$\sup_{x \geq 0} \left( \frac{\sigma x}{1 + \sigma x} - \frac{\eta x}{1 + \eta x} \right) \leq \frac{\sigma - \eta}{(\sqrt{\sigma} + \sqrt{\eta})^2}, \quad \forall 0 \leq \eta \leq \sigma. \quad (21)$$

Also, note that  $\frac{\eta x}{1 + \eta x}$  has derivative bounded above by  $\eta$  on  $x \geq 0$ , and  $x - \kappa_m(x) \leq 1/m$ . Thus, by considering separately the cases  $x \leq m$  and  $x \geq m$

$$\sup_{x \geq 0} \left( \frac{\eta x}{1 + \eta x} - \frac{\eta \kappa_m(x)}{1 + \eta \kappa_m(x)} \right) \leq \max \left\{ \frac{\eta}{m}, \frac{1}{1 + \eta m} \right\}, \quad \forall \eta \geq 0, \forall m \geq 1.$$

Therefore, given  $\lambda > \lambda_c$  and  $\varepsilon > 0$  we may choose  $\eta < \sigma(\lambda)$  such that for all integers  $m$  large enough

$$\sup_{x \geq 0} \left( \frac{\sigma(\lambda)x}{1 + \sigma(\lambda)x} - \frac{\eta \kappa_m(x)}{1 + \eta \kappa_m(x)} \right) < \varepsilon. \quad (22)$$

For this choice of  $\eta$ , choose  $m$  large enough so that  $\lambda > \lambda_c^{(m)}$ ,  $\eta < \sigma^{(m)}(\lambda)$ , and (22) holds. Then, there exists a constant  $c > 0$  such that the conclusion of Corollary 2.4 holds for this  $\eta$  and the modified vertex sequence  $\mathbf{w}^{(m)}$ . Then, the monotonicity property (8), the choice of  $\eta$  and  $m$  satisfying (22), and Corollary 2.4 imply

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \inf_{t \leq e^{cn}} \inf_{i \in [n]} \left( P_{\mathbf{w}}^n(\eta_t(i) = 1) - \frac{\sigma(\lambda)w_i}{1 + \sigma(\lambda)w_i} \right) \\ & \geq \liminf_{n \rightarrow \infty} \inf_{t \leq e^{cn}} \inf_{i \in [n]} \left( P_{\mathbf{w}^{(m)}}^n(\eta_t(i) = 1) - \frac{\eta \kappa_m(w_i)}{1 + \eta \kappa_m(w_i)} - \varepsilon \right) \geq -\varepsilon. \end{aligned}$$

$\square$

#### 4. COMPARISON WITH A MULTI-TYPE BRANCHING PROCESS

In this section we introduce a related multi-type branching process (MTBP) that we will use to show that the contact process dies out quickly if  $\lambda < \lambda_c$ . For fixed vertex weights  $\mathbf{w}$  and  $n \geq 1$ , let  $\mathbf{Z}(t) = (Z_1(t), Z_2(t), \dots, Z_n(t))$  be a MTBP with  $n$  types. Each individual dies at rate 1, and individuals of type  $i$  give birth to individuals of type  $j$  at rate  $\lambda w_i w_j / n$ . The relevance of the MTBP  $\mathbf{Z}(t)$  to the contact process  $\eta_t$  is given by the following lemma.

**Lemma 4.1.** *The MTBP  $\mathbf{Z}(t)$  stochastically dominates the contact process  $\eta_t$ . That is,  $\mathbf{Z}(t)$  and  $\eta_t$  can be coupled in such a way that  $\eta_0 = \mathbf{Z}(0)$  and  $\eta_t \leq \mathbf{Z}(t)$  for all  $t \geq 0$ .*

*Proof.* We describe the coupling informally as follows. We will divide the population of the branching process into two groups, one of which will be identified with the contact process  $\eta_t$ . At any time  $t$ , we will denote the individuals in groups 1 and 2 by  $\mathbf{Z}^{(1)}(t)$  and  $\mathbf{Z}^{(2)}(t)$ , respectively, so that  $\mathbf{Z}(0) = \mathbf{Z}^{(1)}(0) + \mathbf{Z}^{(2)}(0)$ . Let  $\mathbf{Z}^{(1)}(0) = \eta_0$ , so that initially all individuals are in group 1. The dynamics of the processes are described as follows. All individuals of type  $i$  in either population die at rate 1 and give birth to an individual of type  $j$  at rate  $\lambda w_i w_j / n$ . An offspring of type  $j$  at time  $t$  is placed in group 1 if and only if the parent was in group 1 and  $Z_j^{(1)}(t) = 0$  (that is, there are no individuals of type  $j$  in group 1 present already). All other offspring are placed in group 2. It is easy to see that  $\mathbf{Z}^{(1)}(t)$  has the same distribution as the contact process  $\eta_t$ . Since  $\mathbf{Z}^{(2)}(t) \geq \mathbf{0}$ , the proof is complete.  $\square$

Based on the above coupling, we will use  $P_{\mathbf{w}}^{n,i}$  and  $P_{\mathbf{w}}^n$  to denote the law of  $\mathbf{Z}(t)$  under the initial conditions  $\mathbf{Z}(0) = \mathbf{e}_i$  and  $\mathbf{Z}(0) = \mathbf{1}$ , respectively. Part (i) of Theorem 1.2 follows immediately from Lemma 4.1 and the following Theorem.

**Theorem 4.2.** *Let  $\lambda < \lambda_c$ . Then there exists a constant  $C > 0$  such that*

$$\lim_{n \rightarrow \infty} P_{\mathbf{w}}^n(\mathbf{Z}(C \log n) = \mathbf{0}) = 1, \quad \mu - a.s.$$

*Proof.* Define the mean matrix  $\mathbf{M}(t) = (M_{i,j}(t))_{i,j=1}^n$  for the MTBP  $\mathbf{Z}(t)$  by

$$M_{i,j}(t) = E_{\mathbf{w}}^{n,i}[Z_j(t)].$$

Then,

$$\frac{dM_{i,j}(t)}{dt} = -M_{i,j}(t) + \sum_{k=1}^n M_{i,k}(t) \frac{\lambda w_k w_j}{n},$$

which in matrix form is

$$\frac{d\mathbf{M}(t)}{dt} = -\mathbf{M}(t)\mathbf{A},$$

where  $\mathbf{A} = (A_{i,j})_{i,j=1}^n$  is the matrix with entries

$$A_{i,j} = \frac{\lambda w_i w_j}{n} - \delta_{i,j}.$$

Solving this system of differential equations gives that

$$\mathbf{M}(t) = e^{\mathbf{A}t}.$$

A MTBP such as  $\mathbf{Z}(t)$  with mean matrix  $\mathbf{M}(t) = e^{\mathbf{A}t}$  is called *sub-critical* if the largest eigenvalue of  $\mathbf{A}$  is negative. It is easy to see from the definition of  $\mathbf{A}$  above that  $\mathbf{A} = \frac{\lambda}{n} \mathbf{w}_n^* \mathbf{w}_n - I$ , where  $\mathbf{w}_n = (w_1, w_2, \dots, w_n)$  is the row vector of weights and  $\mathbf{w}_n^*$  is the transpose of  $\mathbf{w}_n$ . Now,  $\mathbf{w}_n^* \mathbf{w}_n$  is a rank one matrix with  $n-1$  eigenvalues at 0 and one eigenvalue at  $\|\mathbf{w}_n\|^2 = \sum_{j=1}^n w_j^2$ . (To see this, check that  $\mathbf{w}_n$  is an eigenvector with this as the corresponding eigenvalue.) Therefore, the matrix  $\mathbf{A}$  has  $n-1$  eigenvalues at  $-1$  and one eigenvalue at  $-1 + \frac{\lambda}{n} \|\mathbf{w}_n\|^2$ , and the MTBP  $\mathbf{Z}(t)$  is sub-critical if

$$\lambda < \frac{n}{\|\mathbf{w}_n\|^2} = \left( \frac{1}{n} \sum_{j=1}^n w_j^2 \right)^{-1}.$$

For any  $n$  by  $n$  matrix  $\mathbf{Q}$ , let  $\|\mathbf{Q}\| := \sup\{\|\mathbf{Q}\mathbf{v}\| : \|\mathbf{v}\| \leq 1\}$  be the spectral radius of  $\mathbf{Q}$ . Then, the above analysis of the eigenvalues of  $\mathbf{A}$  implies that

$$\|\mathbf{M}(t)\| = e^{-t(1 - \frac{\lambda}{n} \|\mathbf{w}_n\|^2)}.$$

The law of large numbers and the definition of  $\lambda_c$  imply that

$$\lim_{n \rightarrow \infty} \frac{\|\mathbf{w}_n\|^2}{n} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n w_i^2 = \frac{1}{\lambda_c}, \quad \mu - a.s.$$

Thus, if  $\lambda < \lambda_c$  there exists an  $\varepsilon > 0$  such that  $1 - \frac{\lambda}{n} \|\mathbf{w}_n\|^2 \geq \varepsilon$  for all  $n$  large enough,  $\mu - a.s.$  Therefore,  $\mu - a.s.$ ,

$$P_{\mathbf{w}}^{n,i}(Z_j(t) \geq 1) \leq E_{\mathbf{w}}^{n,i}[Z_j(t)] = M_{i,j}(t) \leq \|\mathbf{M}(t)\| \leq e^{-\varepsilon t}, \quad \text{for all } n \text{ large enough.}$$

(The second to last inequality follows from the fact that the entries of  $\mathbf{M}(t)$  are non-negative.) Therefore, even with the initial configuration of  $\mathbf{Z}(0) = \mathbf{1}$  we have that,  $\mu - a.s.$ ,

$$P_{\mathbf{w}}^n(\mathbf{Z}(t) \neq \mathbf{0}) \leq \sum_{i,j=1}^n P_{\mathbf{w}}^{n,i}(Z_j(t) \geq 1) \leq n^2 e^{-\varepsilon t}, \quad \text{for all } n \text{ large enough.}$$

Letting  $t = C \log n$  for some  $C > 2/\varepsilon$  completes the proof of the Theorem.  $\square$

We can also use the branching process approach to give an upper bound on the probability of a vertex to be infected after a long time. The following Proposition complements the lower bound in Corollary 3.1 and thus completes the proof of Theorem 1.3.

**Proposition 4.3.** *Let  $\lambda > \lambda_c$ . Then, there exists a  $C > 0$  such that*

$$\limsup_{n \rightarrow \infty} \sup_{t \geq C \log n} \sup_{i \in [n]} \left( P_{\mathbf{w}}^n(\eta_t(i) = 1) - \frac{\sigma(\lambda)w_i}{1 + \sigma(\lambda)w_i} \right) \leq 0.$$

*Proof.* The self-duality property (9) and Lemma 4.1 imply that

$$P_{\mathbf{w}}^n(\eta_t(i) = 1) = P_{\mathbf{w}}^{n,i}(\eta_t \neq \mathbf{0}) \leq P_{\mathbf{w}}^{n,i}(\mathbf{Z}(t) \neq \mathbf{0}). \quad (23)$$

To approximate  $P_{\mathbf{w}}^{n,i}(\mathbf{Z}(t) \neq \mathbf{0})$  we need the following lemma.

**Lemma 4.4.** *For  $\lambda > \left(\frac{1}{n} \sum_{j=1}^n w_j^2\right)^{-1}$ , let*

$$\hat{\sigma}_n(\lambda) \text{ be the unique } \sigma > 0 \text{ that solves} \quad 1 = \frac{\lambda}{n} \sum_{j=1}^n \frac{w_j^2}{1 + \sigma w_j}.$$

*Then, the extinction probability  $\hat{\rho}_{i,n} = \hat{\rho}_{i,n}(\lambda)$  satisfies*

$$\hat{\rho}_{i,n} = \lim_{t \rightarrow \infty} P_{\mathbf{w}}^{n,i}(\mathbf{Z}(t) = \mathbf{0}) = \frac{1}{1 + \hat{\sigma}_n(\lambda)w_i}.$$

*Remark 4.5.* If  $\lambda > \lambda_c$ , then  $\hat{\sigma}_n(\lambda)$  and  $\hat{\rho}_{i,n}$  are defined for all  $n$  large enough. Moreover,

$$\lim_{n \rightarrow \infty} \hat{\sigma}_n(\lambda) = \sigma(\lambda) \quad \text{and} \quad \lim_{n \rightarrow \infty} \hat{\rho}_{i,n} = \rho_i := \frac{1}{1 + \sigma(\lambda)w_i}, \quad \mu - a.s. \quad (24)$$

*Proof.* An equation determining the extinction probabilities for multi-type branching processes is given in [AN72, Section 7.5]. In order to use this, we first need to introduce some notation. Let

$$a_i = 1 + \sum_{j=1}^n \frac{\lambda w_i w_j}{n}, \quad f_i(\mathbf{s}) = \frac{1}{a_i} + \sum_{j=1}^n \frac{\lambda w_i w_j}{a_i n} s_i s_j, \quad (25)$$

and

$$\mathbf{u}(\mathbf{s}) = (u_1(\mathbf{s}), u_2(\mathbf{s}), \dots, u_n(\mathbf{s})), \quad \text{where} \quad u_i(\mathbf{s}) = a_i(f_i(\mathbf{s}) - s_i). \quad (26)$$

Then, the extinction probability vector  $\hat{\rho}_n = (\hat{\rho}_{1,n}, \hat{\rho}_{2,n}, \dots, \hat{\rho}_{n,n})$  is the unique solution to

$$\mathbf{u}(\mathbf{s}) = \mathbf{0}, \quad \mathbf{0} \leq \mathbf{s} < \mathbf{1}. \quad (27)$$

Let  $\sigma > 0$  and  $\mathbf{s} = (s_1, s_2, \dots, s_n)$ , where

$$s_i = \frac{1}{1 + \sigma w_i}$$

Recalling (25) and (26) we obtain that

$$\begin{aligned} u_i(\mathbf{s}) &= 1 + \frac{\lambda}{n} w_i s_i \sum_{j=1}^n w_j s_j - s_i - \frac{\lambda}{n} w_i s_i \sum_{j=1}^n w_j \\ &= (1 - s_i) - \frac{\lambda}{n} w_i s_i \sum_{j=1}^n w_j (1 - s_j) \\ &= \frac{\sigma w_i}{1 + \sigma w_i} - \frac{\lambda}{n} \frac{w_i}{1 + \sigma w_i} \sum_{j=1}^n \frac{\sigma w_j^2}{1 + \sigma w_j} \\ &= \frac{\sigma w_i}{1 + \sigma w_i} \left( 1 - \frac{\lambda}{n} \sum_{j=1}^n \frac{w_j^2}{1 + \sigma w_j} \right). \end{aligned}$$

The proof is completed by noting that the term inside the parenthesis on the last line above equals zero when  $\sigma = \hat{\sigma}_n(\lambda)$ .  $\square$

Recall that by (23) we need an upper bound for  $P_{\mathbf{w}}^{i,n}(\mathbf{Z}(t) \neq \mathbf{0})$ . By Lemma 4.4, we know that for  $n$  fixed  $P_{\mathbf{w}}^{i,n}(\mathbf{Z}(t) \neq \mathbf{0})$  decreases to  $1 - \rho_{i,n}$  as  $t$  increases. We would like to show that this convergence is fast enough so that when  $t$  is large enough the error is very small.

To this end, let  $T_0 := \inf\{t > 0 : \mathbf{Z}(t) = \mathbf{0}\}$  be the extinction time of the MTBP  $\mathbf{Z}(t)$ . Let  $\bar{P}_{\mathbf{w}}^{n,i}(\cdot) = P_{\mathbf{w}}^{n,i}(\cdot | T_0 < \infty)$  be the law of  $\mathbf{Z}(t)$  started from one individual of type  $i$  and conditioned to eventually die out. Recall that if  $\lambda > \lambda_c$  then for all  $n$  large enough  $\hat{\rho}_{i,n} = P_{\mathbf{w}}^{n,i}(T_0 < \infty) \in (0, 1)$ . Then,

$$\begin{aligned} P_{\mathbf{w}}^{n,i}(\mathbf{Z}(t) \neq \mathbf{0}) &= P_{\mathbf{w}}^{n,i}(\mathbf{Z}(t) \neq \mathbf{0}, T_0 < \infty) + P_{\mathbf{w}}^{n,i}(\mathbf{Z}(t) \neq \mathbf{0}, T_0 = \infty) \\ &\leq P_{\mathbf{w}}^{n,i}(T_0 < \infty) \bar{P}_{\mathbf{w}}^{n,i}(\mathbf{Z}(t) \neq \mathbf{0}) + P_{\mathbf{w}}^{n,i}(T_0 = \infty) \\ &= \hat{\rho}_{i,n} \bar{P}_{\mathbf{w}}^{n,i}(\mathbf{Z}(t) \neq \mathbf{0}) + 1 - \hat{\rho}_{i,n}. \end{aligned} \tag{28}$$

**Lemma 4.6.** *If  $\lambda > \lambda_c$  then there exists an  $\varepsilon > 0$  so that,  $\mu$ -a.s., for all  $n$  large enough*

$$\bar{P}_{\mathbf{w}}^{n,i}(Z_j(t) \neq 0) \leq e^{-\varepsilon t}, \quad \forall i, j \leq n.$$

*Proof.* The MTBP  $\mathbf{Z}(t)$  conditioned to go extinct can be analyzed using  $h$ -transforms. Given the vertex weights  $\mathbf{w}$ , let  $h_n$  be defined by

$$h_n(\mathbf{z}) = \prod_{i=1}^n \hat{\rho}_{i,n}^{z_i}, \quad \mathbf{z} = (z_1, z_2, \dots, z_n).$$

Then the jump rates of the process  $\mathbf{Z}(t)$  conditioned on extinction are given by  $\bar{q}_n(\mathbf{x}, \mathbf{y}) = q_n(\mathbf{x}, \mathbf{y}) h_n(\mathbf{y}) / h_n(\mathbf{x})$ , where  $q_n(\mathbf{x}, \mathbf{y})$  are the jump rates of the process  $\mathbf{Z}(t)$  without conditioning on extinction. Therefore, we obtain that under  $\bar{P}_{\mathbf{w}}^{n,i}$ ,  $\mathbf{Z}(t)$  is a MTBP with modified birth and death rates. Under the measure  $\bar{P}_{\mathbf{w}}^{n,i}$ , individuals of type  $i$  die at rate  $1/\hat{\rho}_{i,n}$  and give birth to particles of type  $j$  at rate  $\frac{\lambda w_i w_j}{n} \hat{\rho}_{j,n}$ . We define the mean matrix conditioned on extinction by  $\bar{\mathbf{M}}(t) = (\bar{M}_{i,j}(t))_{i,j=1}^n$ , where

$$\bar{M}_{i,j}(t) = \bar{E}_{\mathbf{w}}^{n,i}[Z_j(t)],$$

and  $\overline{E}_{\mathbf{w}}^{n,i}$  denotes expectations with respect to  $\overline{P}_{\mathbf{w}}^{n,i}$ . Then, as was done for the mean matrix  $\overline{\mathbf{M}}(t)$  above we see that

$$\overline{\mathbf{M}}(t) = e^{\overline{\mathbf{A}}t}, \quad \text{where} \quad \overline{\mathbf{A}} = (\overline{A}_{i,j})_{i,j=1}^n, \quad \overline{A}_{i,j} = \begin{cases} -\frac{1}{\hat{\rho}_{j,n}} + \frac{\lambda w_i w_j}{n} \hat{\rho}_{j,n} & i = j \\ \frac{\lambda w_i w_j}{n} \hat{\rho}_{j,n} & i \neq j. \end{cases}$$

Letting  $D_{\hat{\rho}_n}$  be the diagonal matrix with diagonal entries  $\hat{\rho}_{i,n}$  we obtain that

$$\overline{\mathbf{A}} = \frac{\lambda}{n} \mathbf{w}_n \mathbf{w}_n^* D_{\hat{\rho}_n} - D_{\hat{\rho}_n}^{-1} = D_{\hat{\rho}_n}^{-1} \left( \frac{\lambda}{n} D_{\hat{\rho}_n} \mathbf{w}_n \mathbf{w}_n^* D_{\hat{\rho}_n} - I \right)$$

As above, we see that the eigen values of  $\frac{\lambda}{n} D_{\hat{\rho}_n} \mathbf{w}_n \mathbf{w}_n^* D_{\hat{\rho}_n} - I$  are  $-1$  (with multiplicity  $n-1$ ) and  $\frac{\lambda}{n} \|D_{\hat{\rho}_n} \mathbf{w}_n\|^2 - 1 = \frac{1}{n} \sum_{i=1}^n w_i^2 \hat{\rho}_{i,n}^2 - 1$ . Note that the definitions of  $\hat{\rho}_{i,n}$  and  $\hat{\sigma}_n(\lambda)$  imply that

$$\frac{1}{n} \sum_{i=1}^n w_i^2 \hat{\rho}_{i,n} = \frac{1}{n} \sum_{i=1}^n \frac{w_i^2}{1 + \hat{\sigma}_n(\lambda)} = 1.$$

Therefore, since  $\hat{\rho}_{i,n} < 1$  we obtain that (for  $n$  large enough)  $\frac{1}{n} \sum_{i=1}^n w_i^2 \hat{\rho}_{i,n}^2 < 1$ , and thus all the eigenvalues of  $\frac{\lambda}{n} D_{\hat{\rho}_n} \mathbf{w}_n \mathbf{w}_n^* D_{\hat{\rho}_n} - I$  are negative. We wish, however, to show that all the eigenvalues of  $\overline{\mathbf{A}}$  are negative. To this end, note that  $D_{\hat{\rho}_n}^{-1}$  is a diagonal matrix with diagonal entries all greater than one. We prove the following algebra lemma in the appendix.

**Lemma 4.7.** *Let  $\mathbf{U}$  be a symmetric matrix with all eigenvalues less than  $\lambda_0 < 0$ , and let  $D$  be a diagonal matrix with diagonal entries all greater than 1. Then,  $D\mathbf{U}$  also has all eigenvalues less than  $\lambda_0$ .*

Applying Lemma 4.7, we obtain that  $\overline{\mathbf{A}}$  has all eigenvalues less than  $\frac{1}{n} \sum_{i=1}^n w_i^2 \hat{\rho}_{i,n}^2 - 1 < 0$ . Recalling the definition of  $\rho_i$  in (24), we obtain that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n w_i^2 \hat{\rho}_{i,n}^2 - 1 = E_{\mu} [w_1^2 \rho_1^2] - 1 = E_{\mu} \left[ \frac{w_1^2}{(1 + \sigma(\lambda) w_1)^2} \right] - 1 < 0,$$

and therefore,  $\mu - a.s.$ , for any  $0 < \varepsilon < 1 - E_{\mu} [w_1^2 / (1 + \sigma(\lambda) w_1)^2]$  and all  $n$  large enough,  $\|\overline{\mathbf{M}}(t)\| \leq e^{-\varepsilon t}$ . The proof is completed by noting that

$$\overline{P}_{\mathbf{w}}^{n,i}(Z_j(t) \neq 0) \leq \overline{M}_{i,j}(t) \leq \|\overline{\mathbf{M}}(t)\|.$$

□

We are now ready to finish the proof of Proposition 4.3. Recalling (23) and (28), Lemmas 4.4 and 4.6 imply that,  $\mu - a.s.$ , for all  $n$  large enough,

$$P_{\mathbf{w}}^n(\eta_t(i) = 1) \leq n e^{-\varepsilon t} + 1 - \hat{\rho}_{i,n} = n e^{-\varepsilon t} + \frac{\hat{\sigma}_n(\lambda) w_i}{1 + \hat{\sigma}_n(\lambda) w_i}, \quad \forall i \leq n.$$

Therefore, recalling (23) we obtain that

$$\begin{aligned} & \sup_{t \geq C \log n} \sup_{i \in [n]} \left( P_{\mathbf{w}}^n(\eta_t(i) = 1) - \frac{\sigma(\lambda) w_i}{1 + \sigma(\lambda) w_i} \right) \\ & \leq \sup_{t \geq C \log n} \sup_{i \in [n]} \left( n e^{-\varepsilon t} + \frac{\hat{\sigma}_n(\lambda) w_i}{1 + \hat{\sigma}_n(\lambda) w_i} - \frac{\sigma(\lambda) w_i}{1 + \sigma(\lambda) w_i} \right) \\ & \leq n^{1-C\varepsilon} + \frac{|\hat{\sigma}_n(\lambda) - \sigma(\lambda)|}{\left( \sqrt{\hat{\sigma}_n(\lambda)} + \sqrt{\sigma(\lambda)} \right)^2}, \end{aligned}$$

where we used (21) in the last inequality. Since  $\hat{\sigma}_n(\lambda) \rightarrow \sigma(\lambda)$  as  $n \rightarrow \infty$ , the conclusion of Proposition 4.3 holds for any  $C > 1/\varepsilon$ . □



5. ASYMPTOTICS OF  $\sigma(\lambda)$ 

In this section we give the proof of Proposition 1.6. The proof doesn't involve probability at all and is purely an exercise in analysis.

Recall that  $\sigma(\lambda)$  is defined for  $\lambda > \lambda_c$  by

$$\sigma(\lambda) \text{ is the unique } \sigma > 0 \text{ that solves } 1 = \lambda E_\mu \left[ \frac{w_1^2}{1 + \sigma w_1} \right].$$

Thus, it is crucial to understand the asymptotics of  $E_\mu \left[ \frac{w_1^2}{1 + \sigma w_1} \right]$  as  $\sigma \rightarrow 0^+$ .

**Lemma 5.1.** *If  $\mu(w_1 \geq x) \sim Cx^{-(\alpha-1)}$  for some  $C > 0$  and  $\alpha > 2$ , then*

$$E_\mu \left[ \frac{w_1^2}{1 + \sigma w_1} \right] \sim \begin{cases} C \frac{(\alpha-1)\pi}{\sin(\pi\alpha)} \sigma^{\alpha-3} & \alpha \in (2, 3) \\ -2C \log \sigma & \alpha = 3, \end{cases} \quad \text{as } \sigma \rightarrow 0^+, \quad (29)$$

and

$$E_\mu \left[ \frac{w_1^2}{1 + \sigma w_1} \right] - \frac{1}{\lambda_c} \sim \begin{cases} C \frac{(\alpha-1)\pi}{\sin(\pi\alpha)} \sigma^{\alpha-3} & \alpha \in (3, 4) \\ 3C \sigma \log \sigma & \alpha = 4 \\ -E_\mu[w_1^3] \sigma & \alpha > 4, \end{cases} \quad \text{as } \sigma \rightarrow 0^+. \quad (30)$$

Moreover, the conclusion in the case  $\alpha > 4$  also holds under the assumption that  $E_\mu[w_1^3] < \infty$ .

*Proof.* We first consider the case when  $\alpha \leq 3$ . Note that in this case  $E_\mu[w_1^2] = \infty$ , and thus  $E_\mu[w_1^2/(1 + \sigma w_1)] \rightarrow \infty$  as  $\sigma \rightarrow 0^+$ . The assumption on the tail decay of  $\mu$  implies that for any  $\varepsilon > 0$ , there exists an  $x_0 = x_0(\varepsilon) < \infty$  such that

$$(C - \varepsilon)x^{-(\alpha-1)} \leq \mu(w_1 \geq x) \leq (C + \varepsilon)x^{-(\alpha-1)}, \quad \forall x \geq x_0.$$

Therefore,

$$\begin{aligned} \int_{x_0}^{\infty} \frac{x^2}{1 + \sigma x} ((C - \varepsilon)(\alpha - 1)x^{-\alpha}) dx &\leq E_\mu \left[ \frac{w_1^2}{1 + \sigma w_1} \right] \\ &\leq \frac{x_0^2}{1 + \sigma x_0} + \int_{x_0}^{\infty} \frac{x^2}{1 + \sigma x} ((C + \varepsilon)(\alpha - 1)x^{-\alpha}) dx. \end{aligned}$$

Note that  $\frac{x_0^2}{1 + \sigma x_0}$  and  $\int_1^{x_0} \frac{x^{2-\alpha}}{1 + \sigma x} dx$  both are bounded as  $\sigma \rightarrow 0^+$ . Therefore it is enough to show

$$\int_1^{\infty} \frac{x^{2-\alpha}}{1 + \sigma x} dx \sim \begin{cases} \frac{\pi}{\sin(\pi\alpha)} \sigma^{\alpha-3} & \alpha \in (2, 3) \\ -\log \sigma & \alpha = 3, \end{cases} \quad \text{as } \sigma \rightarrow 0^+. \quad (31)$$

When  $\alpha = 3$ , we can explicitly evaluate this integral to be  $\log(1 + 1/\sigma)$ , which is asymptotic to  $-\log \sigma$  as  $\sigma \rightarrow 0^+$ . To handle the case when  $\alpha \in (2, 3)$  we first make the change of variables  $z = \sigma x$  in the integral in (31) to obtain

$$\int_1^{\infty} \frac{x^{2-\alpha}}{1 + \sigma x} dx = \sigma^{\alpha-3} \int_{\sigma}^{\infty} \frac{z^{2-\alpha}}{1 + z} dz.$$

However, taking  $\sigma \rightarrow 0^+$  in the limit of the last integral we obtain that for any  $\alpha \in (2, 3)$

$$\lim_{\sigma \rightarrow 0^+} \int_{\sigma}^{\infty} \frac{z^{2-\alpha}}{1 + z} dz = \int_0^{\infty} \frac{z^{2-\alpha}}{1 + z} dz = \frac{\pi}{\sin(\pi\alpha)},$$

where the last equality is a standard exercise in complex analysis. This completes the proof of (31) and thus also the proof of (29).

Next, we consider the case when  $\alpha > 3$ . In this case  $E_\mu[w_1^2] = 1/\lambda_c < \infty$ , and thus

$$E_\mu \left[ \frac{w_1^2}{1 + \sigma w_1} \right] = E_\mu \left[ w_1^2 \left( 1 - \frac{\sigma w_1}{1 + \sigma w_1} \right) \right] = \frac{1}{\lambda_c} - \sigma E_\mu \left[ \frac{w_1^3}{1 + \sigma w_1} \right]. \quad (32)$$

Thus, we are reduced to computing the asymptotics of  $E_\mu[w_1^3/(1 + \sigma w_1)]$ . Note that  $E_\mu[w_1^3/(1 + \sigma w_1)] \rightarrow E_\mu[w_1^3]$  as  $\sigma \rightarrow 0^+$ . This completes the proof of (30) in the case when  $E_\mu[w_1^3] < \infty$  (in particular when  $\alpha > 4$ ). On the other hand, if  $\alpha \leq 4$  then  $E_\mu[w_1^3] = \infty$  and we must compute the asymptotics of how fast  $E_\mu[w_1^3/(1 + \sigma w_1)]$  grows as  $\sigma \rightarrow 0^+$ . To this end, we repeat the process above in the case when  $\alpha \in (2, 3]$  to show that

$$E_\mu \left[ \frac{w_1^3}{1 + \sigma w_1} \right] \sim C(\alpha - 1) \int_1^\infty \frac{x^{3-\alpha}}{1 + \sigma x} dx \sim \begin{cases} -C \frac{(\alpha-1)\pi}{\sin(\pi\alpha)} \sigma^{\alpha-4} & \alpha \in (3, 4) \\ -3C \log \sigma & \alpha = 4, \end{cases} \quad \text{as } \sigma \rightarrow 0^+.$$

Combining these asymptotics with (32) completes the proof of (30) when  $\alpha \in (3, 4]$ .  $\square$

With Lemma 5.1 at our disposal we are now ready to give the proof of Proposition 1.6.

*Proof of Proposition 1.6.* First, note that the definition of  $\sigma(\lambda)$  and the asymptotics in Lemma 5.1 imply that  $\sigma(\lambda) \rightarrow 0$  as  $\lambda \rightarrow \lambda_c$ . Then, another application of Lemma 5.1 implies that

$$1 = \lambda E_\mu \left[ \frac{w_1^2}{1 + \sigma(\lambda) w_1} \right] \sim \begin{cases} \lambda C \frac{(\alpha-1)\pi}{\sin(\pi\alpha)} \sigma(\lambda)^{\alpha-3} & \alpha \in (2, 3) \\ -2C \lambda \log \sigma(\lambda) & \alpha = 3, \end{cases} \quad \text{as } \lambda \rightarrow 0^+.$$

The above asymptotics are equivalent to the conclusion of Proposition 1.6 in the cases  $\alpha \in (2, 3)$  and  $\alpha = 3$ . Similarly, when  $\alpha > 3$  the definition of  $\sigma(\lambda)$  and Lemma 5.1 imply that

$$\frac{\lambda - \lambda_c}{\lambda_c} = \lambda \left( \frac{1}{\lambda_c} - E_\mu \left[ \frac{w_1^2}{1 + \sigma(\lambda) w_1} \right] \right) \sim \begin{cases} -\lambda C \frac{(\alpha-1)\pi}{\sin(\pi\alpha)} \sigma(\lambda)^{\alpha-3} & \alpha \in (3, 4) \\ -\lambda 3C \sigma(\lambda) \log \sigma(\lambda) & \alpha = 4 \\ \lambda E_\mu[w_1^3] \sigma(\lambda) & \alpha > 4, \end{cases} \quad \text{as } \lambda \rightarrow \lambda_c^+.$$

The conclusions of Proposition 1.6 follow easily from the above asymptotics in the cases  $\alpha \in (3, 4)$  and  $\alpha > 4$  (or  $E_\mu[w_1^3] < \infty$ ). In the case  $\alpha = 4$ , the conclusion in Proposition 1.6 follows from the above asymptotics and the fact that

$$f(\delta) \log f(\delta) \sim -A\delta \iff f(\delta) \sim A \frac{\delta}{\log(1/\delta)},$$

where the above asymptotics are as  $\delta \rightarrow 0^+$ .  $\square$

#### APPENDIX A. PROOF OF LEMMA 4.7

Lemma 4.7 is a simple consequence (albeit not necessarily immediate) of the Spectral Theorem, but we give the proof here for completeness.

*Proof of Lemma 4.7.* As in the statement of the Lemma, let  $U$  be a real  $n \times n$  symmetric matrix with largest eigenvalue  $\lambda_0 < 0$ , and let  $D$  be an  $n \times n$  diagonal matrix with diagonal entries  $D_{i,i} \geq 1$ . Let  $\langle \cdot, \cdot \rangle_D$  be the inner product on  $\mathbb{R}^n$  defined by  $\langle \mathbf{x}, \mathbf{y} \rangle_D := \mathbf{x}^* D^{-1} \mathbf{y}$ . Since  $U$  is a symmetric matrix, then the linear map  $\mathbf{x} \mapsto DU\mathbf{x}$  is a symmetric linear map with respect to the inner product  $\langle \cdot, \cdot \rangle_D$ . Therefore, by the Spectral Theorem there exists a basis  $\{\mathbf{v}_k\}_{k=1}^n$  of  $\mathbb{R}^n$  consisting of eigenvectors of  $DU$  that are orthogonal with respect to the inner product  $\langle \cdot, \cdot \rangle_D$ . Moreover, by scaling the basis we may assume that  $\|\mathbf{v}_k\| = 1$  for all  $k$ .

Given the basis vectors  $\{\mathbf{v}_k\} = \{(v_{k,1}, v_{k,2}, \dots, v_{k,n})\}$ , let  $c_k = \langle \mathbf{v}_k, \mathbf{v}_k \rangle_D$ . Then, since the diagonal entries of  $D$  are all at least 1 we obtain that

$$c_k = \mathbf{v}_k^* D^{-1} \mathbf{v}_k = \sum_{i=1}^n v_{k,i}^2 D_{i,i}^{-1} \leq \sum_{i=1}^n v_{k,i}^2 = 1.$$

Since the  $\mathbf{v}_k$  are eigenvectors of  $DU$  we have that  $DU\mathbf{v}_k = \lambda_k\mathbf{v}_k$ . Our goal is to show that  $\lambda_k \leq \lambda_0$ . To this end, note that

$$\mathbf{v}_k^* U \mathbf{v}_k = \langle \mathbf{v}_k, DU\mathbf{v}_k \rangle_D = \langle \mathbf{v}_k, \lambda_k \mathbf{v}_k \rangle_D = \lambda_k c_k.$$

However, the well known variational characterisation of the largest eigenvalue gives that

$$\lambda_0 = \sup\{\mathbf{x}^* U \mathbf{x} : \|\mathbf{x}\| = 1\} \geq \mathbf{v}_k^* U \mathbf{v}_k = \lambda_k c_k.$$

Since  $c_k \leq 1$  and  $\lambda_0 < 0$  this implies that  $\lambda_k \leq \lambda_0/c_k \leq \lambda_0$ .  $\square$

## REFERENCES

- [AN72] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.
- [BBCS05] Noam Berger, Christian Borgs, Jennifer T. Chayes, and Amin Saberi. On the spread of viruses on the internet. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 301–310 (electronic), New York, 2005. ACM.
- [CD09] Shirshendu Chatterjee and Rick Durrett. Contact processes on random graphs with power law degree distributions have critical value 0. *Ann. Probab.*, 37(6):2332–2356, 2009.
- [CL03] Fan Chung and Linyuan Lu. The average distance in a random graph with given expected degrees. *Internet Math.*, 1(1):91–113, 2003.
- [Lig99] Thomas M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.
- [MR98] Michael Molloy and Bruce Reed. The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.*, 7(3):295–305, 1998.
- [PSV01a] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic dynamics and endemic states in complex networks. *Phys. Rev. E*, 63(6):066117, May 2001.
- [PSV01b] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86(14):3200–3203, Apr 2001.
- [Rio05] Oliver Riordan. The small giant component in scale-free random graphs. *Combin. Probab. Comput.*, 14(5-6):897–938, 2005.

JONATHON PETERSON, CORNELL UNIVERSITY, DEPARTMENT OF MATHEMATICS, MALOTT HALL, ITHACA, NY 14853, USA

*E-mail address:* `peterson@math.cornell.edu`

*URL:* `http://www.math.cornell.edu/~peterson`